

CLAIMS

1. Nucleotides encoding the full length or part of the nucleocapsid (NP) protein of Newcastle disease virus (NDV).
2. The nucleotides as claimed in claim 1 characterised in that it has the following nucleotide sequence:

5

10	10	20	30	40	50	60
	ATGTCTTCG	TATTCGATGA	ATACGAGCAG	CTCCTCGCTG	CTCAGACTCG	CCCCAATGGA
	70	80	90	100	110	120
10	GCTCACGGAG	GGGGAGAGAG	AGGGAGCACT	TTAAAGAGTTG	AGGTCCCAGT	ATTCACTCTT
	130	140	150	160	170	180
	AACAGTGACG	ATCCAGAAAGA	TAGATGGAAT	TTTGCAGGTAT	TCTGTCTTCG	GATTGCTGTT
	190	200	210	220	230	240
15	AGCGAGGAGC	CCAAACAAAC	GCTCAGGCAA	GGTGGCTCTCA	TATCCCTCCT	GTGCTCCCCAT
	250	260	270	280	290	300
	TCTCAAGTGA	TGAGGAAACCA	TGTTGCCCTT	GCAGGAAAAC	AGAAATGAGGC	TACACTGACT
	310	320	330	340	350	360
20	GTTCTTGAGA	TCGATGGTTT	TACCAAGCAGC	GTGCCTCAGT	TCAACAAACAG	GAGTGGGGTG
	370	380	390	400	410	420
	TCTGAGGAGA	GAGCACAGAG	ATTCACTGGT	ATAGCAGGGT	CTCTCCCTCG	GGCGTGCAGT
	430	440	450	460	470	480
25	AACGGTACTC	CGTTCGTCAC	GGCTGGGGTT	GAAGATGATG	CACCAGAAGA	TATCACTGAT
	490	500	510	520	530	540
	ACTCTGGAAA	GAATCCTGTC	TATCCAGGCT	CAGGTATGGG	TCACAGTAGC	GAAGGCCATG
	550	560	570	580	590	600
30	ACTGCATATG	AGACAGCAGA	TGAGTCGGAA	ACAAGAAAGAA	TCAATAAGTA	CATGCAGCAA
	610	620	630	640	650	660
	GGCAGAGTCC	AGAAGAAAGTA	CATCCCTCCAC	CCTGTATGCA	GGAGTGCAAT	TCAACTCACA
	670	680	690	700	710	720
	ATCAGACATT	CTCTGGCAGT	CCGCATTITTC	TTAGTTACGG	AGCTTAAGAG	AGGCCGCAAT
	730	740	750	760	770	780
	ACGGCAGGTG	GGAGCTCCAC	GTATTACAC	TTAGTAGGGG	ATGTAGACTC	ATACATCAGG
	790	800	810	820	830	840
	AACACCGGAC	TTACTGCATT	CTTCCTTACA	CTCAAAATATG	GAATTAATAC	CAAGACATCA

850 860 870 880 890 900
 GCCCTAGCAC TCAGCAGCCT CACAGGCAT ATCCAAAAGA TGAAGCAGCT CATGCCTTA
 910 920 930 940 950 960
 TATCGGATGA AGGGAGAAAA TGCGCCGTAC ATGACATTGC TAGGTGACAG TGATCAGATG
 5 970 980 990 1000 1010 1020
 AGCTTTGCAC CGGCTGAGTA TGCACAGCCT TATTCTTTG CCATGGGCAT GGCATCAGTC
 1030 1040 1050 1060 1070 1080
 TTAGATAAAG GAACTGGCAA ATACCAATTG GCCAGAGACT TCATGAGCAC ATCATTCTGG
 1090 1100 1110 1120 1130 1140
 10 1150 1160 1170 1180 1190 1200
 AGACTCGGGG TGGAGTATGC TCAGGCTAG GGGAGTAGCA TCAACGAAGA CATGGCTGCT
 GAGCTAAAC TAACCCCGC AGCAAGAAGG GGCCCTGGCAG CTGCTGCCCA ACAGAGTGTCT
 1210 1220 1230 1240 1250 1260
 GAGGAAACTG GCAGCGTGGA TATTCTACT CAACAAGCCG GGGTCTCAC TGGGCTCAGC
 15 1270 1280 1290 1300 1310 1320
 GATGGAGGCC CCCGAGCCTC TCAGGGTGGA TCGAACAAAGT CGCAAGGGCA ACCAGATGCC
 1330 1340 1350 1360 1370 1380
 GGAGATGGGG AGACCCAATT CTTGGATTG ATGAGAGCAG TGGCGAACAG CATGCGAGAA
 20 1390 1400 1410 1420 1430 1440
 GCGCCAAACT CCGCACAGAG CACCAACCCAC CGGGAACCCC CCCCGACTCC CGGGCCATCA
 1450 1460 1470 1480 1490 1500
 CAAGATAACG ACACCGACTG GGGGTATTGA

3. Nucleotides encoding the full length or part of the phosphoprotein (P) of Newcastle disease virus (NDV).

25 4. The nucleotides as claimed in claim 3 characterised in that it has the following nucleotide sequence:

10 20 30 40 50 60
 ATGCCACCT TTACAGATGC GGAGATAGAT GATATATTG AGACCAGTGG AACTGTCATT
 30 70 80 90 100 110 120
 GACAGCATAA TTACGGCCA GGGTAATCA GCAGAGACTG TCGGAAGGAG CGCAATCCCA
 130 140 150 160 170 180
 CAAGGCAGA CCAAAGCGCT GAGCATAGCA TGGGAGAACG ATGGGAGCAT CCAACCATCC
 190 200 210 220 230 240
 ACCAGCCAGG ACAACCCCGA CCAACAGGAT AGACCAGACA AACAGCTATC CACACCTGAG
 35 250 260 270 280 290 300
 CAGGCGACCC CACACAAACAG CTCGCCAGCC ACATCCGGCG AACCGCTCCC CACTCAGGCC

310	320	330	340	350	360
GCAGGTGAGG	CCGGCGACAC	ACAGCTCAAG	ACCGGAGCAA	GCAACTCTCT	TCTGTCATG
370	380	390	400	410	420
CTCGACAAGC	TGAGCAATAA	ACCATCTAA	GCTAAAAAGG	GCCCCATGGTC	GAGTCCCCAG
430	440	450	460	470	480
GAAGGATATC	ATCAAACCTCC	GACCCAACAA	CATGGGGATC	AGCCGAACCG	CGGAAACAGC
490	500	510	520	530	540
CAGGAGAGGC	TGCGGCACCA	AGCCAAGGCC	GCCCCTGGAA	GCCGGGGCAGC	AGACGCGAGC
550	560	570	580	590	600
ACAGCATATC	ATGGACAATG	GAAGGGATCA	CAACTATCAG	CTGGTGCAGC	CCCTCATGTG
610	620	630	640	650	660
CTCCAATCATG	GGCAGAGGCC	AGACAGTACT	CCTGTACCTG	TGGATCATGT	CCAGGCCACCT
670	680	690	700	710	720
GTCGACTTTG	TGCAGGCGAT	GATGACTATG	ATGGAGGCGT	TATCACAGAA	GGTAAGTAAA
730	740	750	760	770	780
GTCGACTATC	AGCTAGACCT	AGTCTTAAAG	CAGACATCCT	CCATCCCTAT	GATGCGGTCT
790	800	810	820	830	840
GAAATCCAAC	AGCTAAAAAC	ATCTGTTGCG	GTCATGGAAG	CTAATTTAGG	CATGATGAAA
850	860	870	880	890	900
ATTCTGGACC	CTGGTTGTGC	TAACATTCA	TCCTTAAGT	ATCTGCGGGC	AGTCGCCCCG
910	920	930	940	950	960
TCCCCACCCAG	TTTTAATTC	AGGCCCGGA	GATCCGTCCC	CCTACAGTGC	ACAAGGGGGT
970	980	990	1000	1010	1020
GAGATGACAC	TCAATAAACT	CTCACAAACCA	GTACAACACC	CTTCCGAGTT	AATTAATCT
1030	1040	1050	1060	1070	1080
GCCACACCGG	GCGGACCTGA	TATGGGAGTC	AAAAAGGACA	CTGTCCGTGC	ATTGATCACCC
1090	1100	1110	1120	1130	1140
TCGCGCCCGA	TGCATCCAAG	CTCCTCAGCT	AAGCTCTGTA	GTAAGCTGGA	TGCAGCCGGG
1150	1160	1170	1180	1190	1200
TCGATTGAAG	AGATCAGAAA	GATCAAGCCC	CTTGCACCTAA	ATGGCTAA..

5. The NP protein coded according to claim 1 or claim 2 characterised in that it has the following amino acid sequence:

1	M	S	S	V	F	D	E	Y	E	Q	L	L	A	A	Q	T	16
	ATG	TCT	TCC	GTA	TTC	GAT	GAA	TAC	GAG	CAG	CTC	CTC	GCT	GCT	CAG	ACT	
35				10			20			30			40				
17	R	P	N	G	A	H	G	G	G	E	R	G	S	T	L	R	32
	CGC	CCC	AAT	GGA	GCT	CAC	GGA	GGA	GGA	GAG	AGA	GGG	AGC	ACT	TTA	AGA	
	50			60			70					90					

	33	V E V P V F T L N S D D P E D R	48
		GTT GAG GTC CCA GTA TTC ACT CTT AAC AGT GAC GAT CCA GAA GAT AGA	
	100	110 120 130 140	
5	49	W N F A V F C L R I A V S E D A	64
		TGG AAT TTT GCG GTA TTC TGT CTT CGG ATT GCT GTT AGC GAG GAC GCC	
	150	160 170 180 190	
	65	N K P L R Q G A L I S L L C S H	80
		AAC AAA CCG CTC AGG CAA GGT GCT CTC ATA TCC CTC CTG TGC TCC CAT	
	200	210 220 230 240	
10	81	S Q V M R N H V A L A G K Q N E	96
		TCT CAA GTG ATG AGG AAC CAT GTT GCC CTT GCA GGA AAA CAG AAT GAG	
	250	260 270 280	
	97	A T L T V L E I D G F T S S V P	112
		GCT ACA CTG ACT GTT CTT GAG ATC GAT GGT TTT ACC AGC AGC AGT GTG CCT	
15		290 300 310 320 330	
	113	Q F N N R S G V S E E R A Q R F	128
		CAG TTC AAC AAC AGG AGT GGG GTG TCT GAG GAG AGA GCA CAG AGA TTC	
	340	350 360 370 380	
20	129	M V I A G S L P R A C S N G T P	144
		ATG GTG ATA GCA GGG TCT CTC CCT CGG GCG TGC AGT AAC GGT ACT CGG	
	390	400 410 420 430	
	145	F V T A G V E D D A P E D I T D	160
		TTC GTC ACG GCT GGG GTT GAA GAT GAT GCA CCA GAA GAT ATC ACT GAT	
	440	450 460 470 480	
25	161	T L E R I L S I Q A Q V W V T V	176
		ACT CTG GAA AGA ATC CTG TCT ATC CAG GCT CAG GTA TGG GTC ACA GTA	
	490	500 510 520	
	177	A K A M T A Y E T A D E S E T R	192
		GCG AAG GCC ATG ACT GCA TAT GAG ACA GCA GAT GAG TCG GAA ACA AGA	
30		530 540 550 560 570	
	193	R I N K Y M Q Q G R V Q K K Y I	208
		AGA ATC AAT AAG TAC ATG CAG CAA GGC AGA GTC CAG AAG TAC ATC	
	580	590 600 610 620	
35	209	L H P V C R S A I Q L T I R H S	224
		CTC CAC CCT GTA TGC AGG AGT GCA ATT CAA CTC ACA ATC AGA CAT TCT	
	630	640 650 660 670	
	225	L A V R I F L V S E L K R G R N	240
		CTG GCA GTC CGC ATT TTC TTA GTT AGC GRG CTT AAG AGA GGC CGC ATAT	
	680	690 700 710 720	
40	241	T A G G S S T Y N L V G D V D	256
		ACG GCA GGT GGG AGC TCC ACG TAT TAC AAC TTA GTC GGG GAT GTC GAC	
	730	740 750 760	
	257	S Y I R N T G L T A F F L T L K	272
		TCA TAC ATC AGG AAC ACC GGA CTT ACT GCA AAC TTC CTT ACA CTC AAA	
45		770 780 790 800 810	
	273	Y G I N T K T S A L A L S S L T	288
		TAT GGA ATT AAT ACC AAG ACA TCA GCC CTA GCA CTC AGC AGC CTC ACA	
	820	830 840 850 860	
50	289	G D I Q K M K Q L M R L Y R M K	304
		GGC GAT ATC CAA AAG ATG AAG CAG CTC ATG CGT TTA TAT CGG ATG AAG	
	870	880 890 900 910	

	305	G E N A P Y M T L L G D S D Q M	320
		GGA GAA AAT GCG CCG TAC ATG ACA TTG CTA GGT GAC AGT GAT CAG ATG	
	920	930 940 950 960	
5	321	S F A P A E Y A Q L Y S F A M G	336
		AGC TTT GCA CCG GCT GAG TAT GCA CAG CTT TAT TCT TTT GCC ATG GGC	
	970	980 990 1000	
10	337	M A S V L D K G T G K Y Q F A R	352
		ATG GCA TCA GTC TTA GAT AAA GGA ACT GGC AAA TAC CAA TTC GCC AGA	
	1010	1020 1030 1040 1050	
15	353	D F M S T S F W R L G V E Y A Q	368
		GAC TTC ATG AGC ACA TCA TTC TGG AGA CTC CGG GTG GAG TAT GCT CAG	
	1060	1070 1080 1090 1100	
20	369	A Q G S S I N E D M A A E L K L	384
		GCT CAG GGG AGT AGC ATC AAC GAA GAC ATG GCT GAG CTA AAA CTA	
	1110	1120 1130 1140 1150	
25	385	T P A A R R G L A A A Q R V S	400
		ACC CGG GCA GCA AGA AGG GGC CTG GCA GCT GCT GCC CAA CGA GTG TCT	
	1160	1170 1180 1190 1200	
30	401	E E T G S V D I P T Q Q A G V L	416
		GAG GAA ACT GGC AGC GTG GAT ATT CCT ACT CAA CAA GCC GGG GTC CTC	
	1210	1220 1230 1240	
35	417	T G L S D G G P R A S Q G G S N	432
		ACT GGG CTC AGC GAT GGA GGC CCC CGA GCC TCT CAG GGT GGA TCG AAC	
	1250	1260 1270 1280 1290	
40	433	K S Q G Q P D A G D G E T Q F L	448
		AAG TCG CAA GGG CAA CCA GAT GCC GGA GAT GGG GAG ACC CAA TTC TTG	
	1300	1310 1320 1330 1340	
45	449	D L M R A V A N S M R E A P N S	464
		GAT TTG ATG AGA GCA GTG GGG AAC AGC ATG CGA GAA GCG CCA AAC TCC	
	1350	1360 1370 1380 1390	
50	465	A Q S T T H P E P P P T P G P S	480
		GCA CAG AGC ACC ACC CAC CCG GAA CCC CCC CCG ACT CCC GGG CCA TCC	
	1400	1410 1420 1430 1440	
55	481	Q D N D T D W G Y *	490
		CAA GAT AAC GAC ACC GAC TGG GGG TAT TGA	
	1450	1460 1470	

6. The P protein coded according to claim 3 or claim 4 characterised in that it has the following amino acid sequence:

40	1	M A T F T D A E I D D I F E T S	16
		ATG GCC ACC TTT ACA GAT GCG GAG ATA GAT GAT ATA TTT GAG ACC AGT	
	1	10 20 30 40	
45	17	G T V I D S I I T A Q G K S A E	32
		GGG ACT GTC ATT GAC AGC ATA ATT ACG GCC CAG GGT AAA TCA GCA GAG	
	50	60 70 80 90	

	33	T V G R S A I P Q G K T K A L S	48
		ACT GTC GGA AGG AGC GCA ATC CCA CAA GGC AAG ACC AAA CGG CTG AGC	
	100	110 120 130 140	
5	49	I A W E K H G S I Q P S T S Q D	64
		ATA GCA TGG GAG AGG CAT GGG AGC ATC CAA CCA TCC ACC AGC CAG GAC	
	150	160 170 180 190	
	65	N P D Q Q D R P D K Q L S T P E	80
		AAC CCC GAC CAA CAG GAT AGA CCA GAC AAA CAG CTA TCC ACA CCT GAG	
	200	210 220 230 240	
10	81	Q A T P H N S S P A T S A E P L	96
		CAG GCG ACC CCA CAC AAC AGC TCG CCA GCC ACA TCC GCC GAA CCG CTC	
	250	260 270 280	
15	97	P T Q A A G E A G D T Q L K T G	112
		CCC ACT CAG GCC GCA GGT GAG GCC GGC GGC ACA CAG CTC AAG ACC GGA	
	290	300 310 320 330	
	113	A S N S L L S M L D K L S N K F	128
		GCA AGC AAC TCT CTT CTG TCT ATG CTC GAC AAG CTG AGC AAT AAA CCA	
	340	350 360 370 380	
20	129	S N A K K G P W S S P Q E G Y H	144
		TCT AAT GCT AAA AGG GGC CCA TGG TCG AGT CCC CAG GAA GGA TAT CAT	
	390	400 410 420 430	
	145	Q P P T Q Q H G D Q P N R G N S	160
		CAA CCT CCG ACC CAA CAA CAT GGG GAT CAG CCG AAC CGC GGA AAC AGC	
	440	450 460 470 480	
25	161	Q E R L R H Q A K A A P G S R G	176
		CAG GAG AGG CTG CGG CAC CAA GCC AAG GCC CCT GGA AGC CGG GGC	
	490	500 510 520	
	177	T D A S T A Y H G Q W K E S Q L	192
30		ACA GAC GCG AGC ACA GCA TAT CAT GGA CAA TGG AAG GAG TCA CAA CTA	
	530	540 550 560 570	
	193	S A G A T P H V L Q S G Q S Q D	208
		TCA GCT GGT GCA ACC CCT CAT GTG CTC CAA TCA GGG CAG AGC CAA GAC	
	580	590 600 610 620	
35	209	S T P V P V D H V Q P P V D F V	224
		AGT ACT CCT GTC CCT GTG GAT CAT GTC CAG CCA CCT GTC GAC TTT GTG	
	630	640 650 660 670	
	225	Q A M M T M M E A L S Q K V S K	240
		CAG GCG ATG ATG ACT ATG ATG GAG GCG GTC TTA TCA CAG AGG GTA AGT AAA	
	680	690 700 710 720	
40	241	V D Y Q L D L V L K Q T S S I P	256
		GTC GAC TAT CAG CTA GAC CTA GTC TTA AAG CAG ACA TCC TCC ATC CCT	
	730	740 750 760	
	257	M M R S E I Q Q L K T S V A V M	272
45		ATG ATG CGG TCT GAA ATC CAA CAG CTA AAA ACA TCT GTT GCG GTC ATG	
	770	780 790 800 810	
	273	E A N L G M M K I L D P G C A N	288
		GAA GCT AAT TTA GGC ATG ATG AAA ATT CTG GAC CCT GGT TGT GCT AAC	
	820	830 840 850 860	
50	289	I S S L S D L R A V A R S H P V	304
		ATT TCA TCC TTA AGT GAT CTG CGG GCA GTC GCC CGG TCC CAC CCA GTT	
	870	880 890 900 910	

305	L	I	S	G	P	G	D	P	S	P	Y	V	T	Q	G	G	320	
	TTA	ATT	TCA	GGC	CCC	GGA	GAT	CCG	TCC	CCC	TAC	GTG	ACA	CAA	GGG	GST		
	920					930			940			950			960			
5	321	E	M	T	L	N	K	L	S	Q	P	V	Q	H	P	S	E	336
	GAG	ATG	ACA	CTC	AAT	AAA	CTC	TCA	CAA	CCA	GTA	CAA	CAC	CCT	TCC	GAG		
	970				980				990				1000					
10	337	L	I	K	S	A	T	A	G	G	P	D	M	G	V	E	K	352
	TTA	ATT	AAA	TCT	GCC	ACA	GCG	GGC	GGA	CCT	GAT	ATG	GGA	GTG	GAA	AAG		
	1010			1020				1030			1040			1050				
15	353	D	T	V	R	A	L	I	T	S	R	P	M	H	P	S	S	368
	GAC	ACT	GTC	CGT	GCA	TTG	ATC	ACC	TCG	CGC	CCG	ATG	CAT	CCA	AGC	TCC		
	1060			1070				1080			1090			1100				
20	369	S	A	K	L	L	S	K	L	D	A	A	G	S	I	E	E	384
	TCA	GCT	AAG	CTC	CTG	AGT	AAG	CTG	GAT	GCA	GCC	GGG	TCG	ATT	GAA	GAG		
	1110			1120				1130			1140			1150				
25	385	I	R	K	I	K	R	L	A	L	N	G	*					396
	ATC	AGA	AAG	ATC	AAG	CGC	CTT	GCA	CTA	AAT	GGC	TAA						
	1160			1170				1180										

7. A recombinant expression plasmid containing the NDV nucleocapsid gene as claimed in claim 1 or claim 2.
8. A recombinant expression plasmid containing the NDV phosphoprotein gene as claimed in claim 3 or claim 4.
9. The recombinant expression plasmid according to claim 7 which is the expression plasmid pTrcHis2-NP constructed by cloning the NDV nucleocapsid gene of claims 1 or 2 into vector pTrcHis2.
10. The recombinant expression plasmid according to claim 8 which is the expression plasmid pTrcHis2-P constructed by cloning the NDV phosphoprotein gene of claims 3 or 4 into vector pTrcHis2.
11. A transformed *Escherichia coli* with the recombinant expression plasmid according to claim 7 or claim 9.
12. A transformed *Escherichia coli* with the recombinant expression plasmid according to claim 8 or claim 10.

13. The transformed microorganism according to claim 11, which is the transformed *E. coli* TOP10 (pTrcHis2-NP) produced by introducing the recombinant expression plasmid of claim 7 or claim 9 into *E. coli* TOP10.

5 14. The transformed microorganism according to claim 12, which is the transformed *E. coli* (pTrcHis2-P) produced by introducing the recombinant expression plasmid of claim 8 or claim 10 into *E. coli* TOP 10.

15. A fused or non-fused form of NDV nucleocapsid protein isolated and purified from culture of the transformed microorganism of claim 11 or claim 13 characterised in that it has the following amino acid sequence:

10	1	M S S V F D E Y E Q L L A A Q T	16
	ATG TCT TCC GTA TTC GAT GAA TAC GAG CAG CTC CTC GCT GCT CAG ACT		
	1 10 20 30 40		
15	17 R P N G A H G G G E R G S T L R	32	
	CGC CCC AAT GGA GCT CAC GGA GGG GGA GAG AGA GGG AGC ACT TTA AGA		
	50 60 70 80 90		
	33 V E V P V F T L N S D D P E D R	48	
	GTT GAG GTC CCA GTA TTC ACT CTT AAC AGT GAC GAT CCA GAA GAT AGA		
	100 110 120 130 140		
20	49 W N F A V F C L R I A V S E D A	64	
	TGG AAT TTT GCG GTA TTC TGT CTT CGG ATT GCT GTT AGC GAG GAC GCC		
	150 160 170 180 190		
	65 N K P L R Q G A L I S L L C S H	80	
	ACG AAA CCG CTC AGG CAA GGT GCT CTC ATA TCC CTC CTG TGC TCC CAT		
	200 210 220 230 240		
25	81 S Q V M R N H V A L A G K Q N E	96	
	TCT CAA GTG ATG AGG AAC CAT GTT GCC CTT GCA GGA AAA CAG AAC GAT GAG		
	250 260 270 280		
30	97 A T L T V L E I D G F T S S V P	112	
	GCT ACA CTG ACT GTT CTT GAG ATC GAT GGT TTT ACC AGC AGC GTG CCT		
	290 300 310 320 330		
	113 Q F N N R S G V S E E R A Q R F	128	
	CAG TTC AAC AAC AGG AGT GGG GTG TCT GAG GAG AGA GCA CAG AGA TTC		
	340 350 360 370 380		
35	129 M V I A G S L P R A C S N G T P	144	
	ATG GTG ATA GCA GGG TCT CTC CCT CGG GCG TGC AGT AAC GGT ACT CCG		
	390 400 410 420 430		
	145 F V T A G V E D D A P E D I T D	160	
	TTC GTC ACG GCT GGG GTT GAA GAT GAT GCA CCA GAA GAT ATC ACT GAT		
	440 450 460 470 480		
40	161 T L E R I L S I Q A Q V W V T V	176	
	ACT CTG GAA AGA ATC CTG TCT ATC CAG GCT CAG GTA TGG GTC ACA GTA		
	490 500 510 520		

	177	A K A M T A Y E T A D E S E T R	192
		GCG AAG GCC ATG ACT GCA 540 TAT GAG ACA GCA GAT GAG TCG GAA ACA AGA	
	530	550	560
	530	570	
5	193	R I N K Y M Q Q G R V Q K K Y I	208
		AGA ATC AAT AAC TAC ATG CAG CAA GGC AGA GTC CAG AAC AGG TAC ATC	
	590	590	600
	590	610	620
	209	L H P V C R S A I Q L T I R H S	224
		CTC CAC CCT GTA TGC AGG ACT GCA ATT CAA CTC ACA ATC AGA CAT TCT	
	630	640	650
	630	660	670
10	225	L A V R I F L V S E L K R G R N	240
		CTG GCA GTC CGC ATT TTC TTA GTT AGC GAG CTT AAC TTA GCA AGA GGC CGC AAT	
	680	690	700
	680	710	720
	241	T A G G S S T Y Y N L V G D V D	256
		ACG GCA GGT GGG AGC TCC ACG TAT TAC AAC TTA GTA GGG GAT GTA GAC	
15		730	740
		750	760
	257	S Y I R N T G L T A F F L T L K	272
		TCA TAC ATC AGG AAC ACC GGA CTT ACT GCA TTC TTC CTT ACA CTC AAA	
	770	780	790
	770	800	810
20	273	Y G I N T K T S A L A L S S L T	288
		TAT GGA ATT AAT ACC AAG ACA TCA GCC CTA GCA CTC AGC AGC CTC ACA	
	820	830	840
	820	850	860
	289	G D I Q K M Q L M R L Y R M K	304
		GGC GAT ATC CAA AAG ATG AAG CAG CTC ATG CGT TTA TAT CGG ATG AAG	
	870	880	890
	870	900	910
25	305	G E N A P Y M T L L G D S D Q M	320
		GGA GAA AAT GCG CCG TAC ATG ACA TTG CTA GGT GAC AGT GAT CAG ATG	
	920	930	940
	920	950	960
	321	S F A P A E Y A Q L Y S F A M G	336
		AGC TTT GCA CCG GCT GAG TAT GCA CAG CTT TAT TCT TTT GCC ATG GGC	
30		970	980
		990	1000
	337	M A S V L D K G T G K Y Q F A R	352
		ATG GCA TCA GTC TTA GAT AAA GGA ACT GGC AAA TAC CAA TTC GCC AGA	
	1010	1020	1030
	1010	1040	1050
35	353	D F M S T S F W R L G V E Y A Q	368
		GAC TTC ATG AGC TCA TCA TTC TGG AGA CTC GGG GTG GAG TAT GCT CAG	
	1060	1070	1080
	1060	1090	1100
	369	A Q G S S I N E D M A A E L K L	384
		GCT CGG GGG AGT AGC ATC AAC GAA GAC ATG GCT GCT GAG CTA AAA CTA	
	1110	1120	1130
	1110	1140	1150
40	385	T P A A R R G L A A Q R V S	400
		ACC CCG GCA GCA AGA AGG GGC CTG GCA GCT GCT GCC CAA CGA GTG TCT	
	1160	1170	1180
	1160	1190	1200
45	401	E E T G S V D I P T Q Q A G V L	416
		GAG GAA ACT GGC AGC GTG GAT ATT CCT ACT CAA CAA GGC GGG GTC CTC	
	1210	1220	1230
	1210	1240	
	417	T G L S D G G P R A S Q G G S N	432
		ACT GGG CTC AGC GAT GGA GGC CCC CGA GCC TCT CAG GGT GGA TCG AAC	
	1250	1260	1270
	1250	1280	1290
50	433	K S Q G Q P D A G D G E T Q F L	448
		AAG TCG CAA GGG CAA CCA GAT GCC GGA GAT GGG GAG ACC CAA TTC TTG	
	1300	1310	1320
	1300	1330	1340

